

Data Science Practices That have helped us and may resonate :-)

Joy Bonaguro | Chief Data Officer | State of CA





This simple diagnostic can help you assess readiness

Step 2: Use data in decision-making

Define your business need and select the right data approach



Performance Management

Suite of tools for selecting, developing, and managing with metrics for existing programs, services, or contracts



Evaluation & Experiments

Suite of tools for assessing the impact of a program or service or testing a new program or service



Advanced Analytics

Suite of tools for exploring business questions, developing new insights, or developing new decision tools

Ongoing Exploratory Data Analysis and Data Development

Suite of activities to explore existing data to inform new efforts and to identify the need and plan for new datasets. This feeds all the other activities.

Step 1: Get your data house in order

Diagnose your data baseline and develop plan to get to Level 3



Level 1: Data Void

You can't answer basic questions about programs and services.



Level 2: Data Fire Drills

You can answer basic or ad hoc questions but only after scrambling to pull the data together.



Level 3: Data on Demand

Your existing data and measures are available on demand and mandatory reports are automated.



Develop process and conceptual tools to sniff out good data science projects



The Truffle Pig Problem: Identifying good data science

data science projects is the single greatest barrier to adoption

You know you have a truffle pig problem if Can you Can you build a Can I have a automate this warehouse? dashboard? process? s1 s3

Solution: The project typology



Some combination

6

Something else...

Project Type: Find the needle in the haystack





Service Change

Service Issue: Difficult to identify targets in a population Data Science Process: Use existing data and predictive modeling to identify targets Service Change: Engage with target subset of population

Result: Department resources are spent where most needed

Service Issue: How can we be more targeted in our audits of the new tax?





Large gross receipts + tiny payroll



Tiny gross receipts + giant payroll





Current Process

- 500 audits annually
- Selected by random sample

Recommended Process

- 500 audits annually
- 50%: random sample
- 50%: target based on flags
- Outreach and education to subset of flags



Beware of Goodhart's law

"When a measure becomes a target, it ceases to be a good measure."

GOODHART'S LAW

WHEN A MEASURE BECOMES A TARGET, IT CEASES TO BE A GOOD MEASURE

IF YOU MEASURE PEOPLE ON

NUMBEROF NAILS MADE

THEN YOU

E

10003 OF MIGHT GET TINY NAILS



RBA can help counter but not eliminate **Goodhart's Law**

sketcholanations

WEIGHT OF

NAILS MADE

A FEW GIANT,



Incorporate ethics into the data science lifecycle

Develop a repeatable ethical process



Incorporate it into the data science lifecycle



Learn more about our lessons learned!



←

6

4

=

P



Thank you! Looking forward to the discussion!

www.govops.ca.gov/caldata/